Rational Functions, Toda Flows, and LR-like Algorithms

Leonid Faybusovich

Department of Mathematics
University of Notre Dame
Mail Distribution Center
Notre Dame, Indiana, 46556-5683

Submitted by Paul A. Fuhrman

ABSTRACT

We reconsider some classical ideas of H. Rutishauser from the point of view of polynomial realization theory. We describe various properties of both Toda-Rutishauser flows and LR-like algorithms using Moser-Rutishauser and Flaschka-Haine maps.

1. INTRODUCTION

The goal of this paper is to explain the basic properties of the Moser-Rutishauser map from the point of view of the polynomial realization theory. The work of H. Rutishauser has proved to be fundamental for many areas of numerical analysis. In particular, he developed [21] his quotient-difference (QD) algorithm for finding poles of various classes of meromorphic functions. Later [20] he showed that if the data related to the QD algorithm are organized in certain tridiagonal matrices, the QD algorithm takes the form of what is now known as the LR algorithm. The work of Rutishauser was based on the classical ideas of J. Hadamard, A. C. Aitken, and many others. See [19] for the history of the subject. It led to the development of so-called QR-like algorithms for solving eigenvalue problems. One of the best accounts of the original ideas of Rutishauser can be found in Chapter 7 of the book [14], which appeared in 1974. It is interesting that at the same time J. Moser [18] proposed to use the same approach for finding analytic solutions to a class of nonlinear dynamical systems called Toda flows. The same type of dynamical systems was discovered much earlier by Rutishauser [22]. See in this connection [25].

655 Avenue of the Americas, New York, NY 10010 0024-3795/94/$7.00

Very recently Toda-Rutishauser flows have found unexpected applications in conformal quantum field theory [27]. They are used for the computation of partition functions in certain matrix models. On the other hand, one of the versions of the QD algorithm was used in [6] for the decoding the classical Goppa codes. The LR algorithm with shifts can be interpreted as a sequence of the so-called dressing transformations for certain completely integrable Hamiltonian systems. In the present paper we compare various viewpoints on Toda-Rutishauser flows and LR-like algorithms. It turns out that the polynomial realization theory developed within the context of control theory provides a natural framework for the description of numerous properties of the above-mentioned dynamical systems (see in this connection also [23]). We feel that in general the realization theory of rational functions developed by control theorists will prove to be useful for understanding the basic properties of dynamical systems arising in noncontrol applications ranging from conformal quantum field theory to decoding the algebraic-geometric codes.

We now briefly describe the content of the paper. In Section 2 we describe the basic properties of orthogonal polynomials related to minimal realizations of rational functions and of corresponding Hankel determinants. Section 3 is devoted to the description of explicit formulas for the solutions to Toda-Rutishauser flows. Here the so-called tau functions [27] are expressed in terms of Hankel determinants. This yields a key connection used in the above-mentioned applications to the conformal quantum field theory. The Moser-Rutishauser and the Flaschka-Haine map provide two alternative methods for the linearization of both Toda-Rutishauser flows and LR-like algorithms with shifts. We describe the image of the set of Jacobi matrices under the Flaschka-Haine map. This yields a complete solution of the inverse spectral problem for this class of matrices (quite similar to the classical symmetric case). Finally, in Section 4 we describe the basic properties of LR-like algorithms and closely related QD algorithms. Here again the Moser-Rutishauser and Flaschka-Haine maps provide the major technical tools. The part of this material related to the Moser-Rutishauser map is classical and goes back to Rutishauser. On the other hand, our discussions with various people show that these connections are not well understood even by many numerical analysts. In the end we briefly describe the role of Hankel determinants in the original Rutishauser approach to the derivation of the quotient-difference algorithm. It seems that original ideas of Rutishauser may still be very useful for the improvement of existing eigenvalue algorithms [9].

2. MINIMAL TRIPLES AND ORTHOGONAL POLYNOMIALS

We denote by $\mathrm{Rat}(n, R)$ the manifold of proper scalar rational functions of fixed McMillan degree n. Suppose that $f \in \mathrm{Rat}(n, R)$ has the following expansion at infinity:

$$f(z)=\sum_{i=0}^{\infty}\frac{h_i}{z^{i+1}}. \tag{2.1}$$

We denote by P_n the vector space of polynomials with real coefficients of degree not greater than n. Introduce a bilinear symmetric form on P_{n-1} :

$$\langle z^i, z^j\rangle = h_{i+j}, \tag{2.2}$$

$0 \le i, j \le n-1$. Let $H_i = \|a_{st}\|$, $a_{st} = h_{s+t-2}$, $1 \le s, t \le i$, be the Hankel matrices related to (2.1). Consider the basis $1, z, \dots, z^{n-1}$ in P_{n-1}. Let $p_0, \dots, p_{n-1}$ be polynomials in P_{n-1} with the following properties: (1) $\text{span}(p_0, \cdots, p_i) = \text{span}(1, z, \cdots, z^i)$, $i = 0, 1, \dots, n-1$; (2) $p_i(z) = z^i +$ lower-degree terms; (3) $\langle p_i, p_j\rangle = 0$ if $i \ne j$. We will call these polynomials (following Gragg [13]) the Lanczos polynomials of the first kind. They exist if and only if $\det H_i \ne 0$, $i = 1, \cdots, n$, and admit the following well-known [14] description.

PROPOSITION 2.1.

$$p_i(z)=\frac{1}{\det H_i}\det\begin{bmatrix} h_0 & h_1 & \dots & h_i \\ h_1 & h_2 & \dots & h_{i+1} \\ \vdots & \vdots & & \vdots \\ h_{i-1} & h_i & \dots & h_{2i-1} \\ 1 & z & \dots & z^i \end{bmatrix}. \tag{2.3}$$

Proof: It is clear from (2.3) that $p_i(z) = z^i +$ lower-degree terms. On the other hand, $\langle z^j, p_i\rangle = 0, j < i$. ■

As is usual in polynomial realization theory [11], introduce the shift operator $S : P_{n-1} \to P_{n-1}$ in the following way. Let $f = p/q$, $q(z) = z^n - \sigma_1 z^{n-1} - \sigma_2 z^{n-2} - \dots - \sigma_n$. By definition

$$Sz^i = z^{i+1}(\operatorname{mod} q). \tag{2.4}$$

More precisely, $Sz^i = z^{i+1}$, $i = 0, 1, \dots, n-2$, $Sz^{n-1} = \sigma_1 z^{n-1} + \sigma_2 z^{n-2} + \dots + \sigma_n$.

PROPOSITION 2.2. *The shift operator S is symmetric relative to the scalar product (2.2).*

Proof: We must prove that $\langle Sz^i, z^j\rangle = \langle z^i, Sz^j\rangle$ for any i, j. The only nontrivial case is where $i < n-1$, $j = n-1$. We have $\langle Sz^i, z^{n-1}\rangle = h_{i+n}$, $\langle z^i, Sz^{n-1}\rangle = \sigma_1 h_{n+i-1} + \ldots + \sigma_n h_i$. The identities $h_{n+i} = \sigma_1 h_{n+i-1} + \ldots + \sigma_n h_i$ easily follow from $f \in \mathrm{Rat}(n, R)$. ■

This result immediately leads to the recurrence relations between Lanczos polynomials. Indeed, by the definition of the shift operator S,

$$Sp_i = p_{i+1} + c_i p_i + \ldots + c_0, \quad i = 0, 1, \ldots, n-2,$$

for some real c_i. Thus $\langle Sp_i, p_j\rangle = c_j \langle p_j, p_j\rangle, j \leq i$. On the other hand, $\langle Sp_i, p_j\rangle = \langle p_i, Sp_j\rangle = 0$ if $j < i-1$. By (2.3)

$$\langle p_i, p_i\rangle = \langle p_i, z^i\rangle = \frac{\det H_{i+1}}{\det H_i} \neq 0. \tag{2.5}$$

Hence $c_j = 0$ if $j < i-1$. If $i = n-1$, we have $Sp_i = c_{n-1}p_{n-1} + \ldots + c_0$ for some real c_i. Thus $\langle p_i, p_i\rangle c_i = \langle Sp_{n-1}, p_i\rangle = \langle p_{n-1}, Sp_i\rangle = 0$ if $i < n-2$. We arrive at the following recurrence relations:

$$Sp_i = p_{i+1} + a_{i+1}p_i + b_i p_{i-1}, \tag{2.6}$$

$i = 0, 1, \ldots, n-1$. Here $p_{-1} = p_n = 0$,

$$b_i = \frac{\langle Sp_i, p_{i-1}\rangle}{\langle p_{i-1}, p_{i-1}\rangle}, \quad a_{i+1} = \frac{\langle Sp_i, p_i\rangle}{\langle p_i, p_i\rangle}. \tag{2.7}$$

It is easy to find the expressions for b_i via Hankel determinants. Indeed, $\langle Sp_i, p_{i-1}\rangle = \langle p_i, Sp_{i-1}\rangle = \langle p_i, z^i\rangle = \langle p_i, p_i\rangle$. Thus

$$b_i = \frac{\langle p_i, p_i\rangle}{\langle p_{i-1}, p_{i-1}\rangle} = \frac{\det H_{i+1} \det H_{i-1}}{(\det H_i)^2}. \tag{2.8}$$

Here we have used (2.5). Let

$$p_i(z) = z^i + d_{i,i+1}z^{i-1} + \ldots + d_{1,i+1}. \tag{2.9}$$

Then $\langle p_i, Sp_i\rangle = \langle p_i, z^{i+1}\rangle + d_{i,i+1}\langle p_i, p_i\rangle$. Further, $z^{i+1} = p_{i+1} - d_{i+1,i+2}z^i +$ lower degree terms. Thus $\langle p_i, z^{i+1}\rangle = -d_{i+1,i+2}\langle p_i, p_i\rangle$. Thus

$$a_i = d_{i,i+1} - d_{i+1,i+2}. \tag{2.10}$$

By (2.3)

$$d_{i+1,i} = -\frac{\det \hat{H}_i}{\det H_i},$$

where

$$\hat{H}_i = \begin{bmatrix} h_0 & h_1 & \dots & h_{i-2} & h_i \\ h_1 & h_2 & \dots & h_{i-1} & h_{i+1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ h_{i-1} & h_i & \dots & h_{2i-3} & h_{2i-1} \end{bmatrix}. \tag{2.11}$$

We finally obtain

$$a_i = \frac{\det \hat{H}_{i+1}}{\det H_{i+1}} - \frac{\det \hat{H}_i}{\det H_i}. \tag{2.12}$$

By (2.6) the matrix

$$A = \begin{bmatrix} a_1 & b_1 & & & 0 \\ 1 & a_2 & b_2 & & \\ & 1 & \ddots & \ddots & \\ & & \ddots & a_{n-1} & b_{n-1} \\ 0 & & & 1 & a_n \end{bmatrix} \tag{2.13}$$

is the matrix of the shift operator S in the basis $p_0, p_1, \dots, p_{n-1}$. By the very definition of the shift operator S we know that in the basis $1, \dots, z^{n-1}$ its matrix has the companion form

$$B = \begin{bmatrix} 0 & 0 & \cdots & 0 & \sigma_n \\ 1 & 0 & & 0 & \sigma_{n-1} \\ 0 & 1 & & 0 & \sigma_{n-2} \\ \vdots & & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & 1 & \sigma_1 \end{bmatrix}. \tag{2.14}$$

Consider the matrix $D = \|d_{ij}\|$, where $d_{ij} = 0$ for $i > j$, $d_{ii} = 1$, and d_{ij} with $i < j$ are defined in (2.9).

PROPOSITION 2.3.

$$A = D^{-1}BD. \tag{2.15}$$

Proof: Let $M : P_n \to P_n$ be the linear operator defined as follows: $Mz^i = p_i$, $i = 0, 1, \ldots, n-1$. Consider the operator $\tilde{M} = M^{-1}SM$. It is clear that the matrix of $\tilde{M}$ in the basis $1, z, \ldots z^{n-1}$ is A. Thus $A = [M]^{-1}[S][M]$, where we denote by $[M]$ the matrix of M in the basis $1, z \ldots, z^{n-1}$. Since $[M] = D$, $[S] = B$, we obtain (2.14). ∎

We now introduce a class of unit upper Hessenberg matrices, $\mathrm{Hess}(n)$. Namely, $U = \|u_{ij}\| \in \mathrm{Hess}(n)$ if and only if $u_{ij} = 0$ for $i > j+1$ and $u_{i+1,i} = 1$, $i = 1, \ldots, n-1$. Suppose that $q(z) = z^n - \sigma_1 z^{n-1} - \cdots - \sigma_n$ is the characteristic polynomial of U.

PROPOSITION 2.4. *For any $U \in \mathrm{Hess}(n)$ there exists a unique unit upper triangular matrix X such that*

$$U = XBX^{-1}, \tag{2.16}$$

where B is defined in (2.14) and $X = [e_1, Ue_1, \ldots, U^{n-1}e_1]$. Here $e_1, \ldots, e_n$ is the canonical basis in R^n.

Proof: Since $U \in \mathrm{Hess}(n)$, it is clear that $e_1, Ue_1, \ldots, U^{n-1}e_1$ form a basis in R^n. Let $S : R^n \to R^n$ be a linear operator which has matrix U in the basis $e_1, \ldots, e_n$. It is clear that B is the matrix of S in the basis $e_1, Ue_1, \ldots, U^{n-1}e_1$. Let $M : R^n \to R^n$ be a linear operator which is defined as follows: $Me_i = U^{i-1}e_1$, $i = 1, \ldots, n$. The operator $M^{-1}SM$ has the matrix B in the basis $e_1, \ldots, e_n$. Thus $B = [M]^{-1}[S][M]$, where we denote by $[M]$ the matrix of the operator M in the basis $e_1, \ldots, e_n$. Thus $B = X^{-1}UX$. Suppose now that $YUY^{-1} = U$ for some unit upper triangular matrix Y. Then $YU^ie_1 = U^iYe_1 = U^ie_1$ for any i. Since $e_1, Ue_1, \ldots, U^{n-1}e_1$ form a basis in R^n, this implies that Y is the identity matrix. ∎

COROLLARY 2.5. *Under the assumptions of Proposition 2.3*

$$D^{-1} = [e_1, Ae_1, \ldots, A^{n-1}e_1]. \tag{2.17}$$

REMARK 2.6. B. Kostant [15] proved a far-reaching Lie-algebraic generalization of Proposition 2.4.

We denote by $\mathrm{Jac}(n)$ the set of tridiagonal matrices of the form (2.13) with all $b_i \neq 0$.

THEOREM 2.7. *Let $f \in \mathrm{Rat}(n, R)$ and have the expansion (2.1) at infinity. Then f admits a realization in the form*

$$f(z) = (e_1, (zI_n - A)^{-1}e_1), \tag{2.18}$$

$A \in \mathrm{Jac}(n)$ if and only if $h_0 = 1$, $\det H_i \neq 0$, $i = 2, \ldots, n$. Here $(x, y) = x_1y_1 + \cdots + x_ny_n$, x, $y \in R^n$. Such a realization is unique.

Proof: Indeed, if $\det H_i \neq 0$, $i = 1, \ldots, n$, then we can construct Lanczos polynomials $p_0, \ldots, p_{n-1}$ using the Gram-Schmidt orthogonalization procedure applied to the basis $1, z, \ldots, z^{n-1}$ relative to the bilinear form $\langle , \rangle$ in P_{n-1}. The recurrence relations (2.6) determine $A \in \mathrm{Jac}(n)$. We claim that (2.18) holds. Indeed, if $S^i p_0 = c_0 p_0 + c_1 p_1 + \cdots + c_i p_i$ for some real c_i, then $A^i e_1 = c_0 e_1 + \cdots + c_i e_i$, since S has the matrix A in the basis $p_0, \ldots, p_{n-1}$. Thus $\langle S^i p_0, p_0 \rangle = c_0 = (A^i e_1, e_1)$ for any i. Here we use the fact that $\langle p_0, p_0 \rangle = h_0 = 1$. It is clear that the triple $(1, S, 1)$ is minimal. Thus the rational function $z \to \langle 1, (zI_n - S)^{-1} 1 \rangle \in \mathrm{Rat}(n, R)$. It is now sufficient to prove that $\langle S^i 1, 1 \rangle = h_i$, $i = 0, 1, \ldots, 2n-1$. But this easily follows by (2.2) and the relations $h_{n+i} = \sigma_n h_i + \sigma_{n-1} h_{i+1} + \ldots + \sigma_1 h_{n+i-1}$. We have constructed a realization of f in the form (2.18).

Suppose that we have another such realization with $A' \in \mathrm{Jac}(n)$. Since both realizations are minimal, there exists $X \in \mathrm{GL}(n, R)$ such that

$$A' = XAX^{-1}, \quad Xe_1 = e_1, \quad (X^T)^{-1} e_1 = e_1. \tag{2.19}$$

We clearly have

$$\begin{aligned} E_i &= \mathrm{span}(e_1, \ldots, e_i) = \mathrm{span}(e_1, Ae_1, \ldots, A^{i-1} e_1) \\ &= \mathrm{span}(e_1, A' e_1, \ldots (A')^{i-1} e_1) \end{aligned}$$

for any i, since $A, A' \in \mathrm{Jac}(n)$. Thus $XE_i = \mathrm{span}(Xe_1, XAe_1, \ldots, XA^{i-1} e_1) = E_i$ because of (2.19). Hence X is upper triangular. Since

$$\begin{aligned} E_i &= \mathrm{span}(e_1, A^T e_1, \ldots, (A^T)^{i-1} e_1) \\ &= \mathrm{span}(e_1, (A')^T e_1, \ldots, \left[(A')^T\right]^{i-1} e_1), \end{aligned}$$

using the same reasoning, we obtain that $(X^T)^{-1}$ is upper triangular. Hence X is a diagonal matrix. Finally, since both A, A' have units below the main diagonal, (2.19) easily implies that $X = I_n$. ∎

REMARK 2.8. The following identities hold:

$$\begin{aligned} \langle p_i, p_i \rangle &= b_i \langle p_{i-1}, p_{i-1} \rangle, \quad i = 1, 2, \ldots, n-1, && (2.20) \\ \mu_i &= \frac{\langle S^i 1, p_{i-1} \rangle}{\langle p_{i-1}, p_{i-1} \rangle} = a_1 + a_2 + \cdots + a_i, && (2.21) \end{aligned}$$

$i = 1, 2, \ldots, n$. Indeed, (2.20) is equivalent to (2.8). Further, $\langle S^i 1, p_{i-1} \rangle = \langle S^{i-1} 1, S p_{i-1} \rangle = \langle S^{i-1} 1, p_i + a_i p_{i-1} + b_{i-1} p_{i-2} \rangle = a_i \langle S^{i-1} 1, p_{i-1} \rangle + b_{i-1} \langle S^{i-1} 1, p_{i-2} \rangle = a_i \langle p_{i-1}, p_{i-1} \rangle + b_{i-1} \langle S^{i-1} 1, p_{i-2} \rangle$. Here we have used (2.6). Hence, $\mu_i = a_i + \mu_{i-1}$ by (2.20). Observe that $\mu_1 = \langle S1, p_0 \rangle / \langle p_0, p_0 \rangle = h_1 = a_1$.

3. TODA-RUTISHAUSER FLOWS

We now briefly discuss the structure of dynamical systems introduced by H. Rutishauser [22]. B. Kostant [16] developed a deep mathematical theory of these systems and their Lie-algebraic generalizations. We want to describe here some recent ideas of H. Flaschka and L. Haine [10]. We will use the notations B^+ (B^-) for the group of invertible upper (lower) triangular matrices, and N^+ (N^-) for the group of unit upper (lower) triangular matrices. We will denote by $L(B^\pm), L(N^\pm)$ the corresponding Lie algebras. It is clear that

$$\mathrm{gl}(n,R) = L(B^-) \oplus L(N^+). \tag{3.1}$$

Denote by $\pi : \mathrm{gl}(n,R) \to L(B^-)$ the projection of $\mathrm{gl}(n,R)$ onto $L(B^-)$ along $L(N^+)$. If $A \in \mathrm{gl}(n,R)$, then $\pi(A)$ is the lower triangular part of A. Consider the following dynamical system on $\mathrm{gl}(n,R)$:

$$\dot{A} = [A, \pi(A^k)], \quad k = 1, \ldots, n-1. \tag{3.2}$$

Given $A \in \mathrm{GL}(n,R)$, let

$$A = LU, \qquad L \in B^-, \quad U \in N^+. \tag{3.3}$$

We will call (3.3) the Bruhat decomposition of A. It exists [24] if and only if the principal minors of A are all nonzero.

PROPOSITION 3.1. *Let $A_0 \in \mathrm{gl}(n,R)$, and*

$$\psi(\vec{t}) = \exp(t_1 A_0 + t_2 A_0^2 + \ldots + t_{n-1} A_0^{n-1}) = L(\vec{t})U(\vec{t}) \tag{3.4}$$

be the Bruhat decomposition ($\vec{t} = (t_1, \ldots t_{n-1})$). Then

$$A(\vec{t}) = L(\vec{t})^{-1} A_0 L(\vec{t})$$

is such that

$$\frac{\partial A}{\partial t_i}(\vec{t}) = [A(\vec{t}), \pi(A(\vec{t})^i)], \tag{3.5}$$

$i = 1, \ldots n-1,\ A(\vec{0}) = A_0$. *Here* $[\,,\,]$ *is the usual matrix commutator.*

Proof: We have from (3.4)

$$\frac{\partial\psi}{\partial t_i}(\vec{t}) = A_0^i L(\vec{t})U(\vec{t}) = \frac{\partial L}{\partial t_i}(\vec{t})U(\vec{t}) + L(\vec{t})\frac{\partial U}{\partial t_i}(\vec{t}).$$

Hence,

$$A(\vec{t})^i = L^{-1}(\vec{t})\frac{\partial L}{\partial t_i}(\vec{t}) + \frac{\partial U}{\partial t_i}(\vec{t})U(\vec{t})^{-1}.$$

Since

$$L^{-1}(\vec{t})\frac{\partial L}{\partial t_i}(\vec{t}) \in L(B^-), \quad \frac{\partial U}{\partial t_i}(\vec{t})U(\vec{t})^{-1} \in L(N^+),$$

we conclude

$$\pi(A(\vec{t})^i) = L^{-1}(\vec{t})\frac{\partial L}{\partial t_i}(\vec{t}). \tag{3.6}$$

On the other hand,

$$\frac{\partial}{\partial t_i}A(\vec{t}) = \frac{\partial}{\partial t_i}\{(L(\vec{t})^{-1}A_0L(\vec{t})\} = [A(\vec{t}), L^{-1}(\vec{t})\frac{\partial L}{\partial t_i}(\vec{t})].$$

Comparing this with (3.6), we arrive at (3.5). ∎

PROPOSITION 3.2. Jac(n), Hess(n) *are invariant manifolds for dynamical systems (3.2).*

Proof: Indeed, using the notation of Proposition 3.1, observe that

$$A(\vec{t}) = U(\vec{t})A_0U(\vec{t})^{-1}. \tag{3.7}$$

If $A_0 \in$ Hess(n), it easily follows from (3.7) that $A(\vec{t}) \in$ Hess(n). On the other hand, if A_0 is an unreduced lower Hessenberg matrix, then $A(\vec{t})$ has the same property. This follows from Proposition 3.1. Now Jac(n) is the intersection of Hess(n) with the set of unreduced lower Hessenberg matrices and hence an invariant manifold too. ∎

Given $A \in$ gl(n, R), denote by f_A the rational function

$$f_A(z) = (e_1, (zI_n - A)^{-1}e_1). \tag{3.8}$$

The map $\Phi : A \to f_A$ will be called the Moser-Rutishauser map and is the most important construction in this paper.

PROPOSITION 3.3. *Let $A(\vec{t})$ satisfy (3.5), $A(\vec{0}) = A_0$. Then*

$$\Phi(A(\vec{t})) = \sum_{i=0}^{\infty} \frac{h_i(\vec{t})}{z^{i+1}},$$

where

$$h_i(\vec{t}) = \frac{(e_1, \psi(\vec{t})A_0^i e_1)}{(e_1, \psi(\vec{t})e_1)}, \tag{3.9}$$

i=0,1,.... Here $\psi(\vec{t})$ is described in (3.4).

REMARK 3.4. Compare this with [3].

Proof: By (3.7) we have $\Phi(A(\vec{t})) = (U^T(\vec{t})e_1, (zI_n - A_0)^{-1}e_1)$. Now $U^T(\vec{t}) = \psi(\vec{t})^T[L(\vec{t})^{-1}]^T$ by (3.4). Hence

$$\Phi(A(\vec{t})) = \frac{(e_1, \psi(\vec{t})(zI_n - A_0)^{-1}e_1)}{l_{11}(\vec{t})}, \tag{3.10}$$

where $L(\vec{t}) = \|l_{ij}(\vec{t})\|$. On the other hand, $(e_1, \psi(\vec{t})e_1) = (L(\vec{t})^T e_1, e_1) = l_{11}(\vec{t})$. Here we use $U(\vec{t}) \in N^+$, $L(\vec{t}) \in B^-$. Hence (3.10) coincides with (3.9). ∎

We now briefly discuss the relationships between tau functions and Hankel determinants. This enables us to obtain the explicit formulas for solutions of Toda-Rutishauser flows. Given $A \in \mathrm{GL}(n, R)$, denote by $\tau_i(A)$ the ith principal minor of A. In other words,

$$\tau_i(A) = (e_1 \wedge e_2 \wedge \cdots \wedge e_i, Ae_1 \wedge Ae_2 \wedge \cdots \wedge Ae_i). \tag{3.11}$$

Here $(u_1 \wedge u_2 \wedge \ldots \wedge u_i, v_1 \wedge v_2 \wedge \ldots \wedge v_i) = \det \|(u_i, v_j)\|$, $u_i, v_i \in R^n$. Suppose that $A_0 \in \mathrm{Jac}(n)$ has the form 2.13 and $\psi(\vec{t})$ is defined in (3.4). Denote by $H_i(\vec{t})$ the corresponding Hankel matrix (see Proposition 3.3).

PROPOSITION 3.5.

$$\tau_i(\psi(\vec{t})) = \frac{\det H_i(\vec{t})\tau_1^i(\psi(\vec{t}))}{b_1^{i-1}b_2^{i-2}\cdots b_{i-1}}, \tag{3.12}$$

$\tau_0(\vec{t}) = 0$. *In particular,*

$$a_i(\vec{t}) = \frac{\partial}{\partial t_1}[\ln \tau_i(\psi(\vec{t})) - \ln \tau_{i-1}(\psi(\vec{t}))], \quad i = 1, 2, \ldots, n, \tag{3.13}$$

$$b_i(\vec{t}) = \frac{\tau_{i+1}(\psi(\vec{t}))\tau_{i-1}(\psi(\vec{t}))}{\tau_i^2(\psi(\vec{t}))}, \quad i = 1, 2, \ldots, n-1. \tag{3.14}$$

Proof: Observe that $e_1 \wedge A_0 e_1 \wedge \cdots \wedge A_0^{i-1} e_1 = e_1 \wedge \ldots \wedge e_i$. On the other hand, $A_0^T e_1 = b_1 e_2$ (mod e_1.) Now by induction $(A_0^T)^i e_1 = b_1 b_2 \cdots b_i e_{i+1}$ $(\mathrm{mod}(e_1, \ldots, e_i))$. In other words, $e_1 \wedge (A_0^T) e_1 \wedge \ldots \wedge (A_0^T)^{i-1} e_1 = \beta_i e_1 \wedge \ldots \wedge e_i$,where $\beta_i = b_1^{i-1} b_2^{i-2} \cdots b_1$. Consequently,

$$\begin{aligned} \tau_i(\psi(\vec{t})) &= \frac{(e_1 \wedge (A_0^T e_1) \wedge \ldots \wedge (A_0^T)^{i-1} e_1, \psi(\vec{t}) e_1 \wedge \ldots \wedge \psi(\vec{t}) A_0^{i-1} e_1)}{\beta_i} \\ &= \frac{\det \|(e_1, \psi(\vec{t}) A_0^{t+s} e_1)\|}{\beta_i}. \end{aligned} \tag{3.15}$$

In particular, $\tau_1(\psi(\vec{t}) = (e_1, \psi(\vec{t}) e_1)$. Comparing this with (3.9), we arrive at (3.12). The formula (3.14) now immediately follows from (2.8). Introduce a matrix $\Gamma_i(\vec{t}) = \|\gamma_{j,k}(\vec{t})\|$, $1 \le j, k \le i$, $\gamma_{jk}(\vec{t}) = (e_1, \psi(\vec{t}) A_0^{j+k-2} e_1)$. It is clear that $\Gamma_i(\vec{t})$ is a Hankel matrix and

$$\frac{\partial \gamma_{jk}}{\partial t_1}(\vec{t}) = \gamma_{jk+1}(\vec{t});$$

see (3.4). Hence,

$$\frac{\partial}{\partial t_1} \det \Gamma_i(\vec{t}) = \det \hat{\Gamma}_i(\vec{t}),$$

or

$$\frac{\partial}{\partial t_1} \ln \tau_i(\vec{t}) = \frac{\det \hat{\Gamma}_i(\vec{t})}{\det \Gamma_i(\vec{t})}.$$

Observe now that $H_i(\vec{t}) = \Gamma_i(\vec{t})/\gamma_{11}(\vec{t})$, $\hat{H}_i(\vec{t}) = \hat{\Gamma}_i(\vec{t})/\gamma_{11}(\vec{t})$. Consequently,

$$\frac{\partial}{\partial t_1}[\ln \tau_i(\vec{t}) - \ln \tau_{i-1}(\vec{t})] = \frac{\det \hat{H}_i(\vec{t})}{\det H_i(\vec{t})} - \frac{\det \hat{H}_{i-1}(\vec{t})}{\det H_{i-1}(\vec{t})},$$

$i = 1, 2, \ldots$. Comparing this with (2.12), we arrive at (3.14). ■

We recall briefly (see e.g., [1] for more details) that the flag manifold $\mathrm{Fl}(n)$ is defined as a set of sequences of vector subspaces $0 \subset V_1 \subset V_2 \subset \cdots \subset V_{n-1} \subset R^n$ such that dim $V_i = i$. The group $\mathrm{GL}(n, R)$ acts naturally on $\mathrm{Fl}(n)$ in the following way. If $V = (V_1, \ldots, V_{n-1}) \in \mathrm{Fl}(n)$, $g \in \mathrm{GL}(n, R)$, then

$$gV = (gV_1, gV_2, \ldots, gV_{n-1}).$$

We denote by E $[E_-]$ the standard flag $(e_1, \mathrm{span}(e_1, e_2), \mathrm{span}(e_1, \ldots, e_{n-1}))$ $[(e_n, \mathrm{span}(e_n, e_{n-1}), \ldots, \mathrm{span}(e_n, e_{n-1}, \ldots, e_2))]$. Observe that $\mathrm{GL}(n, R)$ acts transitively on $\mathrm{Fl}(n)$, and the stabilizer of E in $\mathrm{GL}(n, R)$ is B^+. Thus $\mathrm{Fl}(n) \approx Gl(n, R)/B^+$, which enables us to define on $\mathrm{Fl}(n)$ the structure of an analytic manifold. The flag manifold $\mathrm{Fl}(n)$ admits an important stratification (a cell decomposition) by orbits of the group N^+. The cells of this decomposition can be parametrized by the group of permutations $\Sigma(n)$ of $[1, n]$. Namely, if $\alpha \in \Sigma(n)$, let $P_\alpha : R^n \to R^n$ be such that $P_\alpha(e_i) = e_{\alpha(i)}$. Let $E(\alpha) = P_\alpha E$. Then

$$\mathrm{Fl}(n) = \bigcup_{\alpha \in \Sigma(n)} N^+ E(\alpha), \tag{3.16}$$

and $N^+E(\alpha) \bigcap N^+E(\alpha') = \emptyset$ if $\alpha \neq \alpha'$. The decomposition (3.16) is called the Bruhat decomposition of the flag manifold $\mathrm{Fl}(n)$. The orbit $C = N^+E_-$ is called the big cell of the decomposition (3.16). It is an open dense subset of $\mathrm{Fl}(n)$. Each $V \in C$ admits a unique representation in the form n^+E_-, $n^+ \in N^+$. The map $V \to n^+$ is an analytic map (isomorphism) of C onto N^+.

After this preparation we are ready to describe the Flaschka-Haine construction [10]. Given $(\sigma_1, \sigma_2, \ldots, \sigma_{n-1}) \in R^{n-1}$, denote by $\mathrm{Hess}(n; \sigma)$ the set of upper Hessenberg matrices from $\mathrm{Hess}(n)$ with the characteristic polynomial $z^n - \sigma_1 z^{n-1} - \cdots - \sigma_n$.

PROPOSITION 3.6. *Given $A \in \mathrm{Hess}(n;\sigma)$, let $U(A) = [e_1, Ae_1, \ldots, A^{n-1}e_1] \in N^+$. The map* $\mathrm{Fl} : \mathrm{Hess}(n;\sigma) \to \mathrm{Fl}(n)$, $\mathrm{Fl}(A) = U^{-1}E_-$ *is an analytic isomorphism of* $\mathrm{Hess}(n;\sigma)$ *onto the big cell C of* $\mathrm{Fl}(n)$.

Proof: The map Fl is certainly smooth. If $\mathrm{Fl}(A) = \mathrm{Fl}(A')$, A, $A' \in \mathrm{Hess}(n;\sigma)$, then $U(A) = U(A')$. But then $A = A'$ by Proposition 2.4. The map Fl is onto C. Indeed, if $V = U^{-1}E_- \in C$, $U \in N^+$, then $A = UBU^{-1} \in \mathrm{Hess}(n;\sigma)$ [see (2.14)] and $\mathrm{Fl}(A) = V$ by Proposition 2.4. Finally, the map Fl^{-1} is analytic by the remarks before the Proposition 3.6. ■

The main observation of Flaschka and Haine [10] is that the map Fl linearizes Toda-Rutishauser flows.

PROPOSITION 3.7. *Suppose that $A_0 \in \mathrm{Hess}(n;\sigma)$. Let $A(\vec{t})$ be the solution of (3.2) described in Proposition 3.1. Then*

$$\mathrm{Fl}(A(\vec{t})) = \exp[-(t_1 B + t_2 B^2 + \cdots + t_{n-1} B^{n-1})]\, \mathrm{Fl}(A_0). \tag{3.17}$$

Here B is the companion matrix (2.14).

Proof: Let $A_0 = U(A_0)BU(A_0)^{-1}$, *where* $U(A_0) = [e_1, A_0e_1, \ldots, A_0^{n-1}e_1]$, *and* B is the companion matrix (2.14). By (3.7), $A(\vec{t}) = [U(\vec{t})U(A_0)] B[U(\vec{t})U(A_0)]^{-1}$. Hence $\mathrm{Fl}(A(\vec{t})) = U(A_0)^{-1}U(\vec{t})^{-1}E_-$. By (3.4) $\mathrm{Fl}(A(\vec{t})) = U(A_0)^{-1}\psi(\vec{t})^{-1}E_-$, since $L(\vec{t})E_- = E$. Further,

$$\psi(\vec{t})^{-1} = U(A_0)\exp[-(t_1B + t_2B^2 + \ldots + t_{n-1}B^{n-1})]U(A_0)^{-1}.$$

Thus we arrive at (3.17). ∎

REMARK 3.8. The analogue of the Flaschka-Haine transformation is well known for the case of symmetric Toda flows. But unlike the symmetric case, this transformation is global. It can be thought of as the finite-dimensional analogue of the inverse spectral transform. Solutions of Toda-Rutishauser flows may have finite escape time. This never happens with flows (3.17). The occurrence of finite escape time corresponds to the situation where the flow (3.17) leaves the big cell [10].

We want now to describe the image of $\mathrm{Jac}(n)$ under the Flaschka-Haine map. It is convenient to identify R^n with P_{n-1} using the linear isomorphism $e_i \to z^{i-1}$, $i = 0, 1, \ldots n$. Under this identification B corresponds to the shift operator S.

PROPOSITION 3.9. *Let* $A \in \mathrm{Jac}(n)$. *Under the above identification*

$$\mathrm{Fl}(A) = (p_{n-1}, \mathrm{span}(p_{n-1}, p_{n-2}), \ldots, \mathrm{span}(p_{n-1}, \ldots, p_1)), \tag{3.18}$$

where $p_0 = 1, p_1, \ldots p_{n-1}$ *are Lanczos polynomials corresponding to* A. *In particular,*

$$\mathrm{Fl}(A) = (V_1, V_2, \ldots, V_{n-1}), \tag{3.19}$$

where

$$V_i = \mathrm{span}(p_{n-1}, Sp_{n-1}, \ldots, S^{i-1}p_{n-1}),$$

$i = 1, \ldots, n-1$.

Proof: If $A \in \mathrm{Jac}(n)$, by Proposition 2.3 we know that $U(A)^{-1}$ is the unit upper triangular matrix whose columns are coefficients of corresponding Lanzcos polynomials. Thus $U(A)^{-1}z^i = p_i$, $i = 0, \ldots, n-1$. Under our identifications $E_- = ((z^{n-1}, \mathrm{span}(z^{n-1}, z^{n-2}), \ldots, \mathrm{span}(z^{n-1}, \ldots, z))$. Hence, $U(A)^{-1}E_-$ has the form (3.18). Now using recurrence relations (2.6), we have

$$\begin{aligned}\mathrm{span}(p_{n-1}, p_{n-2}) &= (\mathrm{span}(p_{n-1}, Sp_{n-1}), \ldots, \mathrm{span}(p_{n-1}, \ldots p_1)) \\ &= (p_{n-1}, Sp_{n-1}, \ldots, S^{n-2}p_{n-1}).\end{aligned}$$

This yields (3.19). ∎

Given $X \in \mathrm{gl}(n, R)$ and an X-cyclic vector $x \in R^n$, denote by $V(X; x)$ the flag $(x, \mathrm{span}(x, Xx), \ldots, \mathrm{span}(x, Xx, \ldots, X^{n-1}x))$.

THEOREM 3.10. *The Flaschka-Haine map defines the (analytic) isomorphism of* $\mathrm{Jac}(n;\sigma) = \mathrm{Jac}(n) \bigcap \mathrm{Hess}(n;\sigma)$ *with the set of flags* $V(B;x)$, *where* x *runs through the subset in* R^n *determined by the following conditions. The minors of the matrix*

$$\Gamma(B;x) = [B^{n-1}x, B^{n-2}x, \dots, x] \tag{3.20}$$

obtained by deleting the first i *columns and first* i *rows are all nonzero,* $i = 1, \dots, n-1$, $x_n = 1$. *Here* B *is the companion matrix (2.14).*

Proof: If $A \in \mathrm{Jac}(n;\sigma)$, by Proposition 3.9 $\mathrm{Fl}(A) = V(B;x)$, where $x \in R^n$ is such that $x_1 + x_2 z + \cdots + x_n z^{n-1}$ is the nth Lanzcos polynomial of A. Hence, $x_n = 1$ and $V(B;x) = U(A)^{-1}E_-$. It is clear that $V(B;x) = \Gamma(B;x)E_-$. Hence, $\Gamma(B;x) = U(A)^{-1}L$ for some $L \in B_-$. We see that $\Gamma(B;x)$ admits the Bruhat decomposition. Hence all required minors of this matrix are nonzero. Suppose now that we are given the flag $V(B;x) \in Fl(n)$ such that all conditions are satisfied. Then let

$$\Gamma(B;x) = U^{-1}L, \quad U \in N_+, \quad L \in B_-,$$

be the Bruhat decomposition of $\Gamma(B;x)$. We want to prove that $A = UBU^{-1} \in \mathrm{Jac}(n)$. It is clear that $A \in \mathrm{Hess}(n;\sigma)$. Further,

$$A = L\Gamma(B;x)^{-1}B\Gamma(B;x)L^{-1}.$$

Observe that $\Gamma(B;x)e_i = B^{n-i}x$, $i = 1, 2, \dots, n$. Denote $\Gamma(B;x)^{-1}B\Gamma(B;x)$ by Y. We have $Ye_i = e_{i-1}$, $i = 2, 3, \dots, n$, $Ye_1 = \sigma_1 e_n + \sigma_2 e_{n-1} + \dots + \sigma_n e_1$. Hence Y is a lower Hessenberg matrix with all entries equal to 1 on the first superdiagonal. Thus LYL^{-1} is lower Hessenberg with nonzero entries on the first superdiagonal. Hence $A \in \mathrm{Jac}(n;\sigma)$. ■

PROPOSITION 3.11. *Let* $A \in \mathrm{Jac}(n;\sigma)$, *and* H_n *be the associated Hankel matrix (see (2.18)). Let* $H_n = LU$, $L \in B_-, U \in N^+$, *be the Bruhat decomposition of* H_n. *Then* $U = U(A)$. *In particular,* $U(A)$ *is completely determined by its first row.*

Proof: We know that $U(A) = [e_1, Ae_1, \dots, A^{n-1}e_1]$. Set

$$L(A) = [e_1, A^T e_1, \dots, (A^T)^{n-1} e_1]^T.$$

It is clear that $L(A) \in B^-$ for $A \in \mathrm{Jac}(n)$. Further, $H_n = L(A)U(A)$ is clearly the Bruhat decomposition. Observe that H_n is completely determined by Markov parameters $h_i(A) = (e_1, A^i e_1)$, $i = 0, 1, \dots, n-1$, together with $\sigma_1, \dots, \sigma_n$. But $h_i(A)$, $i = 0, 1, \dots, n-1$, is exactly the first row of $U(A)$. ■

4. QD AND LR ALGORITHMS

In this section we describe various versions of the LR algorithm for finding eigenvalues of matrices. It can be reduced (via the Moser-Rutishauser map) to the power iterations of the shift operator in P_{n-1}. It practically coincides with the progressive form of the quotient-difference (QD) algorithm of Rutishauser. We also show that under the Flashka-Haine transformation the LR algorithm goes to inverse power iterations on the flag manifold. Other versions of the QD algorithm are also discussed.

Let $A \in \mathrm{GL}(n, R)$ be such that all principal minors of A are nonzero. Consider the Bruhat decomposition of A

$$A = LU, \quad L \in B^-, \quad U \in N^+. \tag{4.1}$$

Suppose that $\alpha_0, \alpha_1, \dots$ are real numbers and $A_0 \in \mathrm{gl}(n, R)$ is such that $D_0 = A_0 - \alpha_0 I_n$ admits the Bruhat decomposition (3.3); then if $D_0 = L_0 U_0$, $L_0 \in B^-$, $U_0 \in N^+$, set $A_1 = L_0^{-1} A_0 L_0 = U_0 L_0 + \alpha_0 I_n = U_0 A_0 U_0^{-1}$. Suppose we have already constructed $A_0, \dots, A_i$, and $D_i = A_i - \alpha_i I_n$ admits the factorization (3.3): $D_i = L_i U_i$. Set $A_{i+1} = L_i^{-1} A_i L_i = U_i L_i + \alpha_i I_n = U_i A_i U_i^{-1}$. The described process is known [24, 12] as the LR algorithm with variable shifts $\alpha_0, \alpha_1, \dots$. The key observations [24] related to this algorithm are described as follows. First of all,

$$\begin{aligned} A_{i+1} &= L_i^{-1} A_i L_i = L_i^{-1} L_{i-1}^{-1} A_{i-1} L_{i-1} L_i \cdots \\ &= (L_i^{-1} L_{i-1}^{-1} \cdots L_0^{-1}) A_0 (L_0 L_1 \dots L_i), \end{aligned}$$

or

$$A_{i+1} = (L_0 L_1 \dots L_i)^{-1} A_0 (L_0 L_1 \dots L_i). \tag{4.2}$$

Similarly,

$$A_{i+1} = (U_i U_{i-1} \dots U_0) A_0 (U_i U_{i-1} \dots U_0)^{-1}. \tag{4.3}$$

Denote $L_0 L_1 \dots L_i$ by $\hat{L}_i$ and $U_i U_{i-1} \dots U_0$ by $\hat{U}_i$. By (4.2) we have $\hat{L}_i(A_{i+1} - \alpha_{i+1} I_n) = (A_0 - \alpha_{i+1} I_n)\hat{L}_i$. But $A_{i+1} - \alpha_{i+1} I_n = L_{i+1} U_{i+1}$. Thus we can conclude that

$$\hat{L}_{i+1} U_{i+1} = (A_0 - \alpha_{i+1} I_n)\hat{L}_i. \tag{4.4}$$

If we know that $(A_0 - \alpha_0 I_n) \dots (A_0 - \alpha_i I_n) = \hat{L}_i \hat{U}_i$, we can derive from (4.4) that

$$(A_0 - \alpha_0 I_n) \dots (A_0 - \alpha_{i+1} I_n) = \hat{L}_{i+1} \hat{U}_{i+1}. \tag{4.5}$$

Thus by induction on i we obtain that (4.5) holds true for any i. Recall that the Moser-Rutishauser map was defined as

$$A \to f_A, f_A(z) = (e_1, (zI_n - A)^{-1} e_1). \tag{4.6}$$

PROPOSITION 4.1. *Let A_{i+1} be the $(i+1)$th iteration of the LR algorithm with shifts. Then*

$$f_{A_{i+1}}(z) = \frac{(e_1, (A_0 - \alpha_0 I_n) \cdots (A_0 - \alpha_i I_n)(zI_n - A_0)^{-1} e_1)}{(e_1, (A_0 - \alpha_0 I_n) \ldots (A_0 - \alpha_i I_n) e_1)}, \tag{4.7}$$

$$f_{A_{i+1}}(z) = \frac{(e_1, (A_i - \alpha_i I_n)(zI_n - A_i)^{-1} e_1)}{(e_1, (A_i - \alpha_i I_n) e_1)}. \tag{4.8}$$

Proof: We prove (4.7). The proof of (4.8) is completely similar. We have by (4.2)

$$f_{A_{i+1}}(z) = (\hat{U}_i^T e_1, (zI_n - A_0)^{-1} \hat{U}_i^{-1} e_1).$$

Now $\hat{U}_i^{-1} e_1 = e_1$, since $\hat{U}_i \in N^+$, $\hat{U}_i^T = \varphi_i(A_0)^T (\hat{L}_i^{-1})^T \in B^+$, where $\varphi_i(z) = (z - \alpha_0) \cdots (z - \alpha_i)$. Since $(\hat{L}_i^{-1})^T \in B^+$, we have $\hat{U}_i^T e_1 = \gamma \phi_i(A_0)^T e_1$ for some nonzero γ. This yields

$$f_{A_{i+1}}(z) = \gamma(e_1, \varphi_i(A_0)(zI_n - A_0)^{-1} e_1). \tag{4.9}$$

Let

$$f_A(z) = \sum_{i=0}^{\infty} \frac{h_i(A)}{z^{i+1}} \tag{4.10}$$

be the expansion of f_A at infinity. We clearly have

$$h_i(A) = (e_1, A^i e_1). \tag{4.11}$$

In particular, $h_0(A) = 1$. Applying this to (4.9), we will have

$$h_k(A_{i+1}) = \gamma(e_1, \varphi_i(A) A^k e_1), \tag{4.12}$$

$1 = h_0(A_{i+1}) = \gamma(e_1, \varphi_i(A) e_1)$. ■

It is convenient to introduce (as in [3]) a projectivization PRat(n, R) of Rat(n, R). Namely, we identify two rational functions $f_1, f_2 \in$ Rat(n, R) if $f_1 = \lambda f_2$ for some positive λ. Denote by ρ : Rat$(n, R) \to$ PRat(n, R) the corresponding projection. Let $f = p/q \in$ Rat(n, R), $\deg q = n$, and φ be some polynomial with real coefficients. Suppose that φ and q have no common divisors. Define a map M_φ : PRat$(n, R) \to$ PRat(n, R) in the following way:

$$M_\varphi(\rho(p/q)) = \rho(\tilde{p}/q), \tag{4.13}$$

where $\tilde{p}$ is uniquely defined by the conditions $\deg \tilde{p} < n$, $\tilde{p} = \varphi p \pmod{q}$. In other words, using the shift operator S, $\tilde{p} = \varphi(S)p$. Denote by Rat$_q(n, R)$ the set $\{f = p/q \in$ Rat(n, R) : p, q have no common divisors $\}$.

PROPOSITION 4.2. *Let $q \in R[z]$, $\deg q = n$, be such that $q(\alpha_i) \neq 0$, $i = 0, 1, \ldots$. Then the LR algorithm with variable shifts $\alpha_0, \alpha_1, \ldots$ induces a time-dependent dynamical system D on $P(\mathrm{Rat}_q(R))$ of the form*

$$D(\rho(p/q), i) = M_{z-\alpha_i}(\rho(p/q)) = \rho\left(\frac{(S - \alpha_i I_n)p}{q}\right), \tag{4.14}$$

i=0,1,....

Proof: By (4.8),(4.10),

$$h_k(A_{i+1}) = \frac{(e_1, (A_i - \alpha_i I_n)A_i^k e_1)}{(e_1, (A_i - \alpha_i I_n)e_1)} = \frac{h_{k+1}(A_i) - \alpha_i h_k(A_i)}{h_1(A_i) - \alpha_i h_0(A_i)}. \tag{4.15}$$

Let $f_{A_i} = p/q$ and $\tilde{p}$ be such that $\deg p < n$, $(z - \alpha_i)p = \tilde{p} \pmod q$. It is clear that $(z - \alpha_i)p/q$ differs from $\tilde{p}/q$ by a constant. If $p/q = \sum_{i=0}^{\infty} h_i/z^{i+1}$, then $(z - \alpha_i)p/q = \sum_{i=0}^{\infty}(h_{i+1} - \alpha_i h_i)/z^{i+1} + h_0$. Hence

$$\frac{\tilde{p}}{q} = \sum_{i=0}^{\infty} \frac{h_{i+1} - \alpha_i h_i}{z^{i+1}}. \tag{4.16}$$

Comparing (4.15) with (4.16), we obtain $\rho(f_{A_{i+1}}) = M_{z-\alpha_i}(\rho(p/q))$. ■

EXAMPLE 4.3. This example is given for illustration only. Suppose that $q(z) = (z - \nu_1)\cdots(z - \nu_n)$ is a polynomial with real roots $\nu_1 > \cdots > \nu_n$. Consider the rational function $f = 1/q \in \mathrm{Rat}(n, R)$, and apply to this function the LR algorithm without shifts. According to Proposition 4.2 we should consider the space $P(P_{n-1})$. If $p \in P_{n-1}$, then one iteration of the LR algorithm takes $\rho(p)$ to $\rho(Sp)$. In other words, the $LR-$ algorithm is reduced to the power iterations of the shift operator S. In our case S has eigenvalues $\nu_1, \ldots, \nu_n$ and eigenvectors

$$v_i(z) = \frac{q\gamma_i}{z - \nu_i},$$

where γ_i are chosen in such a way that $v_i(\nu_i) = 1$ (the so-called interpolation basis). Now

$$1 = \sum_{i=1}^{n} v_i.$$

Thus

$$S^k 1 = \sum_{i=1}^{n} \nu_i^k v_i.$$

Hence, $\rho(S^k 1) \to \rho(v_1)$. Observe that $Sv_1 = \nu_1 v_1$. We can now "deflate" the zero ν_1. Consider the rational function $1/v_1$ and apply the LR algorithm again. Observe that v_1 has roots $\nu_2, \ldots, \nu_n$.

REMARK 4.4. Proposition 4.2 shows that the Moser-Rutishauser map linearizes the LR algorithm with variable shifts. The same is true for many other types of QR-like algorithms [7].

Earlier we introduced polynomials $\varphi_i(z) = (z - \alpha_0) \cdots (z - \alpha_i)$, $i = 0, 1, \ldots$. It is clear that $1, \varphi_0, \ldots, \varphi_{i-1}$ form a basis in the space P_i. In general, we have expressions of the form

$$z^i = c_0^{(i)} 1 + c_1^{(i)} \varphi_0 + \ldots + c_{i-1}^{(i)} \varphi_{i-1}. \tag{4.17}$$

PROPOSITION 4.5. *Let $A_0, A_1, \ldots$ be iterations of the LR algorithm with variable shifts $\alpha_0, \alpha_1, \ldots$. Suppose that $A_i - \alpha_i I_n = L_i U_i$, $i = 0, 1, \ldots$, $L_i \in B^-$, $U_i \in N^+$. If*

$$L_i = \|l_{st}^{(i)}\|, \qquad U_i = \|u_{st}^{(i)}\|, \tag{4.18}$$

then

$$h_k(A_0) = (e_1, A_0^k e_1) = c_0^{(k)} + c_1^{(k)} l_{11}^{(0)} + \ldots + c_{k-1}^{(k)} l_{11}^{(0)} l_{11}^{(1)} \cdots l_{11}^{(k-1)}. \tag{4.19}$$

REMARK 4.6. In the case $\alpha_0 = \alpha_1 = \cdots = 0$, the expressions (4.19) are reduced to

$$h_0 = 1, \quad h_k(A_0) = l_{11}^{(0)} \cdots l_{11}^{(k-1)}, \quad k = 1, 2, \ldots. \tag{4.20}$$

In particular,

$$l_{11}^{(k)} = \frac{h_{k+1}(A_0)}{h_k(A_0)}, \quad k = 0, 1, \ldots. \tag{4.21}$$

In general, the equations (4.19) enable us to find $l_{11}^{(k)}$ recursively. First one can find $l_{11}^{(0)}$ from the expression for $h_1(A_0)$, then $l_{11}^{(1)}$ from the expression for $h_2(A_0)$, etc. In [6] we showed how to use Proposition 4.5 for decoding Goppa codes.

Proof: By (4.17) $h_k(A_0) = c_0^{(k)} + c_1^k(e_1, \varphi_0(A_0)e_1) + \ldots + c_{k-1}^{(k)}(e_1, \varphi_{k-1}(A_0)e_1)$. By (4.5) $\varphi_i(A_0) = \hat{L}_i\hat{U}_i$. Hence $(e_1, \varphi_i(A_0)e_1) = (e_1, \hat{L}_i\hat{U}_ie_1) = (\hat{L}_i^Te_1, e_1) = (L_i^TL_{i-1}^T \cdots L_0^Te_1, e_1) = l_{11}^{(0)} \cdots l_{11}^{(i)}$. This yields (4.19). ■

Suppose now that $A \in \mathrm{Jac}(n)$. The Bruhat decomposition of A has a special structure $A = LU$ with

$$L = \begin{bmatrix} l_{11} & 0 & 0 & \cdots & 0 \\ 1 & l_{22} & 0 & \cdots & 0 \\ 0 & 1 & l_{33} & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & l_{nn} \end{bmatrix}, \tag{4.22}$$

$$U = \begin{bmatrix} 1 & u_{12} & 0 & \cdots & 0 & 0 \\ 0 & 1 & u_{23} & \cdots & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & 1 & u_{n-3,n-2} & 0 \\ 0 & \cdots & 0 & 0 & 1 & u_{n-2,n-1} \\ 0 & \cdots & 0 & 0 & 0 & 1 \end{bmatrix}. \tag{4.23}$$

Indeed, $U = L^{-1}A$ is lower Hessenberg (and unit upper triangular), and $L = AU^{-1}$ is in $\mathrm{Hess}(n)$ and is lower triangular. From (4.22), (4.23) we easily obtain

$$a_i = l_{ii} + u_{i-1,i}, \quad i = 1, 2, \ldots, n, \tag{4.24}$$

$$b_i = l_{ii}u_{i,i+1}, \quad i = 1, 2, \ldots, n-1. \tag{4.25}$$

Here $u_{0,1} = 0$. The matrix $\tilde{A} = UL$ will again be in $\mathrm{Jac}(n)$, and

$$\tilde{a}_i = l_{ii} + u_{i,i+1}, \quad i = 1, 2, \ldots, n, \tag{4.26}$$

$$\tilde{b}_i = l_{i+1,i+1}u_{i,i+1}, \quad i = 1, 2 \ldots, n-1. \tag{4.27}$$

Here $u_{n,n+1} = 0$. We can rewrite (4.24)–(4.27) as recursive relations for the LR algorithm with shifts:

$$a_i^{(k)} - \alpha_k = l_{ii}^{(k)} + u_{i-1,i}^{(k)}, \quad i = 1, 2, \ldots, n, \tag{4.28}$$

$$b_i^{(k)} = l_{ii}^{(k)}u_{i,i+1}^{(k)}, \quad i = 1, 2, \ldots, n-1, \tag{4.29}$$

$$a_i^{(k+1)} = l_{ii}^{(k)} + u_{i,i+1}^{(k)} + \alpha_k, \quad i = 1, 2, \ldots, n, \tag{4.30}$$

$$b_i^{(k+1)} = l_{i+1,i+1}^{(k)}u_{i,i+1}^{(k)}, \quad i = 1, 2, \ldots, n-1. \tag{4.31}$$

REMARK 4.7. α_k can be chosen as functions of $a_i^{(k)}, b_i^{(k)}$. For example, if we take $\alpha_k = a_n^{(k)}$, we arrive at the Raileigh rule for the choice of variable shifts.

Suppose,first, that α_k and the entries $l_{11}^{(k)}$, $k = 0, 1, \dots$, are known. Using the fact that $u_{0,1}^k = 0$ and (4.28), we can find $a_1^{(k)} = \alpha_k + l_{11}^{(k)}$, $k = 0, 1, \dots$. Then from (4.30) we can find $u_{12}^{(k)}$, $k = 0, 1, \dots$. We further find $b_1^{(k)}$, $k = 0, 1, \dots$, by (4.29), and then the entries $l_{22}^{(k)}$ can be determined from (4.31) for $k = 0, 1, \dots$. The general recursive scheme can be described as follows:

$$a_i^{(k)} = \alpha_k + l_{ii}^{(k)} + u_{i-1,i}^{(k)}, \tag{4.32}$$

$$u_{i,i+1}^{(k)} = a_i^{(k+1)} - l_{ii}^{(k)} - \alpha_k = a_i^{(k+1)} - a_i^{(k)} + u_{i-1,i}^{(k)}, \tag{4.33}$$

$$b_i^{(k)} = l_{ii}^{(k)} u_{i,i+1}^{(k)}, \tag{4.34}$$

$$l_{i+1,i+1}^{(k)} = b_i^{(k+1)} / u_{i,i+1}^{(k)}. \tag{4.35}$$

Observe that we can find $l_{11}^{(k)}$ using Proposition 4.5. The original scheme of the quotient-difference algorithm coincided with (4.32)– (4.35) (with $\alpha_k = 0$). Although it turned out to be numerically unstable [14], it can be used in an efficient way for the decoding of algebraic-geometric codes [6]. The so-called progressive form of the QD algorithm works as follows. Suppose that we are given initial data $l_{ii}^{(0)}, u_{i,i+1}^{(0)}$ which satisfy (4.28)–(4.29). We can rewrite (4.28)–(4.29) in the form

$$a_i^{(k+1)} - \alpha_{k+1} = l_{ii}^{(k+1)} + u_{i-1,i}^{(k+1)}, \tag{4.36}$$

$$b_i^{(k+1)} = l_{ii}^{(k+1)} u_{i,i+1}^{(k+1)}. \tag{4.37}$$

Combining this with (4.30), (4.31), we obtain

$$l_{ii}^{(k+1)} + u_{i-1,i}^{(k+1)} + \alpha_{k+1} = l_{ii}^{(k)} + u_{i,i+1}^{(k)} + \alpha_k, \tag{4.38}$$

$$l_{ii}^{(k+1)} u_{i,i+1}^{(k+1)} = l_{i+1,i+1}^{(k)} u_{i,i+1}^{(k)}. \tag{4.39}$$

Relations (4.38), (4.39) (with $\alpha_k = 0$) form the "rhombus rules" of Rutishauser. We can rewrite them in the form

$$l_{ii}^{(k+1)} = l_{ii}^{(k)} + \left(u_{i,i+1}^{(k)} - u_{i-1,i}^{(k+1)}\right) + (\alpha_k - \alpha_{k+1}), \tag{4.40}$$

$$u_{i,i+1}^{(k)} = \frac{l_{i+1,i+1}^{(k)} u_{i,i+1}^{(k)}}{l_{ii}^{(k+1)}}. \tag{4.41}$$

The relations (4.40), (4.41) determine the progressive form of the QD algorithm. Observe that in this scheme α_k may depend on $a_i^{(k)}, b_i^{(k)}$. For example, $\alpha_k = l_{n,n}^{(k)} + u_{n-1,n}^{(k)}$ corresponds to the Raileigh shift [24] but the scheme (4.40), (4.41) should be modified in the following way [see (4.28)–(4.31)]:

$$u_{i,i+1}^{(k)} = \frac{b_i^{(k)}}{a_i^{(k)} - \alpha_k - u_{i-1,i}^{(k)}}, \tag{4.42}$$

$$l_{ii}^{(k)} = a_i^{(k)} - \alpha_k - u_{i-1,i}^{(k)}, \tag{4.43}$$

$$a_i^{(k+1)} = l_{ii}^{(k)} + u_{i,i+1}^{(k)} + \alpha_k, \tag{4.44}$$

$$b_i^{(k+1)} = l_{i+1,i+1}^{(k)} u_{i,i+1}^{(k)}. \tag{4.45}$$

PROPOSITION 4.8. *The sets* $\mathrm{Hess}(n), \mathrm{Jac}(n), \mathrm{Hess}(n;\sigma)$ *are invariant manifolds for the LR algorithm with variable shifts.*

Proof: The proof is the same as for Proposition 3.2. ■

PROPOSITION 4.9. *Suppose that* $A_0 \in \mathrm{Hess}(n;\sigma)$ *and* $A_0, A_1, \ldots$ *are iterations of the LR algorithm with variable shifts. Then*

$$\mathrm{Fl}(A_{i+1}) = (B - \alpha_i I_n)^{-1}\mathrm{Fl}(A_i), \quad i = 0, 1, \ldots. \tag{4.46}$$

Here B is defined in (2.14).

Thus the LR algorithm with variable shifts is completely equivalent to power iterations (4.46) on the flag manifold.
Proof: Let $A_i - \alpha_i I_n = L_i U_i$ be the Bruhat decomposition. Then $A_{i+1} = U_i A_i U_i^{-1}$. Let

$$A_i = U_i(A) B U_i^{-1}, \quad U_i(A) \in N^+. \tag{4.47}$$

Then $A_{i+1} = U_i U_i(A) B [U_i U_i(A)]^{-1}$. Hence $\mathrm{Fl}(A_{i+1}) = U_i(A)^{-1} U_i^{-1} E_-$. Using the fact that $A_i - \alpha_i I_n = L_i U_i$, we obtain $\mathrm{Fl}(A_{i+1}) = U_i(A)^{-1} U_i(A)(B - \alpha_i I_n)^{-1} U_i(A)^{-1} E_-$. Here we have used the fact that $L_i^{-1} E_- = E_-$. ■

REMARK 4.10. Relationships between LR-like algorithms and power iterations on flag manifolds are well known. The use of the Flaschka-Haine map is quite different from the standard approach [26, 6, 1] and opens new opportunities for the analysis of convergence properties of LR algorithms.

We want to present another expression for diagonal entries of a Jacobi matrix via Hankel determinants. Let A be a Jacobi matrix of the form (2.13) which admits a Bruhat decomposition $A = LU$ of the form (4.22), (4.23). If $\tilde{A} = UL$ and $\tilde{A} = \tilde{U}\tilde{L}$ is the Bruhat decomposition, we conclude using (4.29), (4.31) that

$$\tilde{b}_i = \frac{l_{i+1,i+1}}{l_{i,i}} b_i, \quad i = 1, \ldots, n-1, \quad l_{11} = h_1. \tag{4.48}$$

Here $h_0, h_1, \ldots$ are Markov parameters of the rational function f_A [see (3.8)]. Hence

$$\tilde{b}_i \tilde{b}_{i-1} \cdots \tilde{b}_1 = \frac{l_{i+1,i+1}}{l_{11}} b_i b_{i-1} \cdots b_1. \tag{4.49}$$

Using (2.8), we see that

$$b_1 \cdots b_i = \frac{\langle p_i, p_i \rangle}{\langle p_0, p_0 \rangle} = \frac{\det H_{i+1}}{\det H_i}.$$

Similarly,

$$\tilde{b}_1 \cdots \tilde{b}_i = \frac{\det \tilde{H}_{i+1}}{\det \tilde{H}_i}.$$

Here $\tilde{H}_i$ are Hankel matrices of the rational function $f_{\tilde{A}}$. Thus by (4.49)

$$l_{i,i} = h_1 \frac{\det \tilde{H}_i \det H_{i-1}}{\det \tilde{H}_{i-1} \det H_i}.$$

We now use (4.15) to conclude that

$$\det \tilde{H}_i = \frac{1}{h_1^i} \det H_i^1, \tag{4.50}$$

where

$$H_i^{(1)} = \begin{bmatrix} h_1 & h_2 & \ldots & h_i \\ h_2 & h_3 & \ldots & h_{i+1} \\ \vdots & \vdots & & \vdots \\ h_i & h_{i+1} & \ldots & h_{2i-1} \end{bmatrix}.$$

Hence

$$l_{ii} = \frac{\det H_i^{(1)} \det H_{i-1}}{\det H_{i-1}^{(1)} \det H_i}. \tag{4.51}$$

By (2.8), (4.29), (4.51), we have

$$u_{i,i+1} = \frac{b_i}{l_{ii}} = \frac{\det H_{i-1} \det H_{i+1} \det H_i \det H_{i-1}^{(1)}}{\det H_i^2 \det H_{i-1} \det H_i^{(1)}} = \frac{\det H_{i+1} \det H_{i-1}^{(1)}}{\det H_i \det H_i^{(1)}}. \tag{4.52}$$

Now by (4.28)

$$a_i = l_{ii} + u_{i-1,i} = \frac{\det H_{i-1} \det H_i^{(1)}}{\det H_{i-1}^{(1)} \det H_i} + \frac{\det H_i \det H_{i-2}^{(1)}}{\det H_{i-1}^{(1)} \det H_{i-1}},$$

$i = 2, 3, \ldots, n,\ a_1 = h_1$. Here $\det H_0^{(1)} = 1$.

REMARK 4.11. We have, of course, similar expressions for $\tilde{l}_{ii}, \tilde{u}_{i,i+1}$. One can now use the "rhombus rules" (4.38), (4.39) to find recurrence relations between Hankel determinants. It is these relations that Rutishauser used [14] to derive the QD algorithm. Here we derive these relations starting from the quotient-difference scheme.

5. CONCLUDING REMARKS

In the present paper we have reconsidered some classical ideas of H. Rutishauser from the point of view of the polynomial realization theory. The emphasis is on the Moser-Rutishauser map (3.8) rather than on continuous fraction expansions and determinant identities. We have described in detail the Flaschka-Haine map, which leads to an alternative way of linearizing Toda-Rutishauser flows and LR-like algorithms.

We strongly believe that profound ideas of Rutishauser will prove to be useful not only in such traditional areas as numerical linear algebra but also in coding theory and in the theory of completely integrable Hamiltonian systems. In particular, the use of the Moser-Rutishauser map leads to the construction of action-angle variables for various classes of such systems [8, 4, 17].

REFERENCES

1 G. Ammar and C. Martin, The geometry of matrix eigenvalue methods, *Acta Appl. Math.* 5:239–278 (1986).

2 M. Chu, On the global convergence of the Toda lattice for real normal matrices and its application to the eigenvalue problem, *SIAM J. Math. Anal.* 15:98–104 (1984).

3 R. W. Brockett and L. Faybusovich, Toda flows, inverse spectral transform and realization theory, *Systems Control Lett.* 16:79–88 (1991).

4 P. Deift, L. C. Li, T. Nanda, and C. Tomei, The Toda flow on a generic orbit is integrable, *Comm. Pure Appl. Math.* 39:183–232. (1986).

5 L. Faybusovich, Toda flows and isospectral manifolds, *Proc. Amer. Math. Soc.* 115:837–845 (1992).

6 L. Faybusovich, On the Rutishauser approach to eigenvalue problems (preprint).

7 L. Faybusovich, Rutishauser approach to eigenvalue problems: Realization theory viewpoint, in *Proceedings of the International Symposium on Implicit and Nonlinear Systems*, 1992, pp. 108–112.

8 L. Faybusovich, On the Hamiltonian and system-theoretic properties of the HR-algorithm, *J. Math. Systems Estim. Control* 3:129–146 (1993).

9 K. Fernando and B. Parlett, Accurate singular values and differential qd algorithms (preprint).

10 H. Flaschka and L. Haine, Variété de drapeaux et réseaux de Toda, *Math. Z.* 208:545–556 (1991).

11 P. Fuhrmann, Functional models in linear algebra, *Linear Algebra Appl.* 162–164:107–151 (1992).

12 G. Golub and C. Van Loan, *Matrix Computations,* Johns-Hopkins U.P., 1989.

13 W. B. Gragg, Matrix interpretations and applications of continued fraction algorithm, *Rocky Mountain J. Math.* 4:213–225 (1974).

14 P. Henrici, *Applied and Computational Complex Analysis*, Vol. 1, Wiley, 1974.

15 B. Kostant, On Whittaker vectors and representation theory, *Invent. Math.* 48:101–184 (1978).

16 B. Kostant, The solution to a generalized Toda lattice and representation theory, *Adv. Math.* 34:195–338 (1979).

17 P. S. Krishnaprasad, Symplectic mechanics and rational functions, *Ricerche Automat.* 10:107–135 (1979).

18 J. Moser, Finitely many points on the line under the influence of exponential potential—an integrable system, in Lecture Notes in Phys. (J. Moser, Ed.), *Dynamical Systems and Applications*, 38:467–497 (1975).

19 B. Parlett, The development and use of methods of LR type, *SIAM Rev.* 6:275–295 (1964).

20 H. Rutishauser, Solution of eigenvalue problems with LR-transformations, *Nat. Bur. Standards Appl. Math. Ser.* 49:47–81 (1958).

21 H. Rutishauser, Der Quotienten-Differenzen-Algorithmus, *Z. Angew. Math. Phy.* 5:233–251 (1954).

22 H. Rutishauser, Ein infinitesimales Analogon zum Quotient-Differenzen-Algorithmus, *Arch. Math. Basel* 5:132–137 (1954).

23 G. Stewart, On a companion operator for analytic functions, *Numer. Math.* 18:26–43 (1971).

24 D. Watkins, *Fundamentals of Matrix Computations*, Wiley, 1991.

25 D. Watkins and L. Elsner, On Rutishauser's approach to self-similar flows,

SIAM J. Matrix Anal. Appl. 11:301–311 (1990).
26 D. Watkins and L. Elsner, Convergence of algorithms of decomposition type for the eigenvalue problem, *Linear Algebra Appl.* 143:19–47 (1991).
27 E. Witten, Two-dimensional gravity and intersection theory on moduli space, in *Surveys in Differential Geometry*, 1991.

Received 10 February 1993; final manuscript accepted 11 July 1993